

## Peer Review of State Assessment Systems Outcomes Report

---

To comply with Section 1111(e) of the Elementary and Secondary Education Act (ESEA) as amended by the Every Student Succeeds Act (ESSA), states are required to submit evidence associated with their assessment systems to the United States Department of Education (ED) for peer review. The peer review process is designed to ensure that states have implemented rigorous content standards and high-quality assessments aligned to those standards. As part of the peer review process, states are required to submit documentation of the technical quality and validity of state assessments, as well as documentation for assessment administration and handling processes.

States must submit documentation for all reading/language arts (R/LA), mathematics, science, and English language proficiency (ELP) assessments that are used for accountability purposes. For each state assessment submitted, the peer review process is conducted by cohorts of trained experts who evaluate state-submitted evidence as it relates to a number of established criteria, organized by and divided into Critical Elements (CEs). Based on this review, ED will issue [decision letters](#) that identify whether states have addressed all CEs and met peer review requirements or must submit additional information for review. Each decision letter identifies the CEs for which states must submit additional evidence to meet the statutory and regulatory requirements for content standards and assessment.

Drawing on an analysis of the state decision letters, this report will look at the CEs for which many states have been required to submit additional evidence across R/LA, mathematics, and science assessments for grades 3–8 and high school; ELP assessments for grades K–12; and alternate assessments. This report highlights peer review responses and identifies areas where multiple states were requested to submit additional information or documentation to meet all peer review requirements.

## Peer Review of State Assessment Systems

The assessment peer review process is in place to ensure that states implement rigorous academic standards and high-quality assessments. The process is evidence-based, focusing on documentation of the process used to develop and administer the assessments, as well as data to confirm the quality of the assessment system. The review is conducted by cohorts who are external assessment experts. All states must submit documentation for peer review of general and alternate assessments in R/LA, mathematics, science, and ELP. The peer review process must be conducted each time a new statewide assessment is implemented<sup>1</sup>.

There are four possible determinations for each component of a state assessment system:

- **Meets Requirements** – The component meets all of the requirements of the statute and regulations. The state is not required to submit any additional documentation for peer review unless significant changes are made to the state’s standards or assessment.
- **Substantially Meets Requirements** – The component meets most of the requirements of the statute and regulations. The state is required to submit some additional information, and the ED expects that this information should be able to be provided within one year. The state must submit a timeline for submission of the additional documentation, and the ED may take additional action if adequate progress is not made.
- **Partially Meets Requirements** – The component does not meet a number of the requirements of the statute and regulations, and/or the state will need to submit substantial information to show that it meets the requirements. The ED expects that this information may not be able to be provided within one year, and a condition will be placed on the state’s Title I grant award. A condition may also potentially be placed upon the state’s grant award under Part B of the Individuals with Disabilities Education Act (IDEA) if the matter pertains to requirements in IDEA. The state will participate in regularly scheduled progress calls with ED, and, depending upon the component in question, the state may be placed on high-risk status and/or the state may be asked to enter into a compliance agreement.
- **Does Not Meet Requirements** – The component does not meet most of the requirements of the statute and regulations and will require substantial work by the state to revise the assessment system design in order to meet the requirements. It is unlikely that the state will be able to make the necessary changes within one year, so the state may be required to halt administration of the assessment component in question. The State will participate in regularly scheduled progress calls with ED, and a condition will be placed on the state’s Title I grant award. A condition may also potentially be placed upon the state’s grant award under Part B of IDEA if the matter pertains to requirements in IDEA. Depending upon the component in question, the state may be placed on high-risk status and/or the state may be asked to enter into a compliance agreement.<sup>2</sup>

## Summary of Peer Review Outcomes

Each decision letter issued to a state provides detailed information on which assessments a state must provide additional evidence for to meet the peer review requirements for a specific CE. This report includes summaries of state determinations, which outline the number of states that met requirements and those that have not yet met requirements, disaggregated by subject.

<sup>1</sup> *A State’s Guide to the U.S. Department of Education’s Assessment Peer Review Process.* (2018). Washington, DC: U.S. Department of Education. Retrieved from <https://www2.ed.gov/admins/lead/account/saa/assessmentpeerreview.pdf>

<sup>2</sup> Whalen, A. (2016, October 6). *Dear colleague.* [Letter from the U.S. Department of Education to the Council of Chief State School Officers, State Assessment Directors, State Title I Directors, and State Special Education Directors]. Retrieved from <https://www2.ed.gov/admins/lead/account/saa/dclletterassepeerreview1072016ltr.pdf>

The following charts (Charts 1–3) provide an overview of the number of states that have received determinations for R/ LA, mathematics, science, and ELP assessments across grade levels, disaggregated by the four possible determinations that states can receive for a submitted assessment: Meets Requirements, Substantially Meets Requirements, Partially Meets Requirements, and Does Not Meet Requirements. These charts provide an overview of the number of states that have met all peer review requirements for content assessments and the number of states that were asked to submit additional evidence.

Not all statewide assessments are included in decision letter feedback, as the assessment may not have been submitted for review or the assessment has not yet received a peer review determination. Charts 1–3 identify the number of states that have not yet received a peer review determination for individual assessments.

Chart 1. Peer Review Determinations for State Reading/Language Arts and Mathematics Assessments

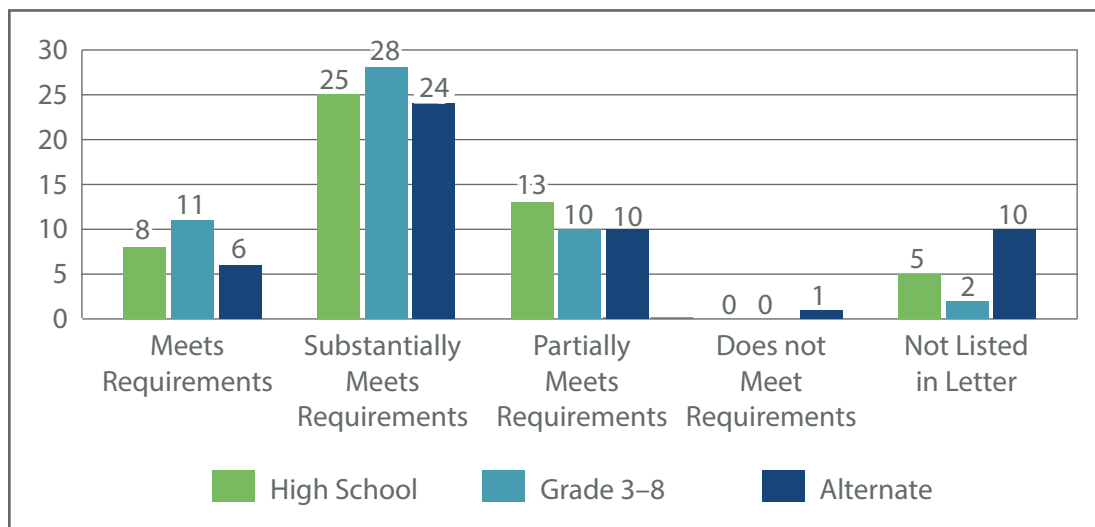


Chart 2. Peer Review Determinations for State Science Assessments

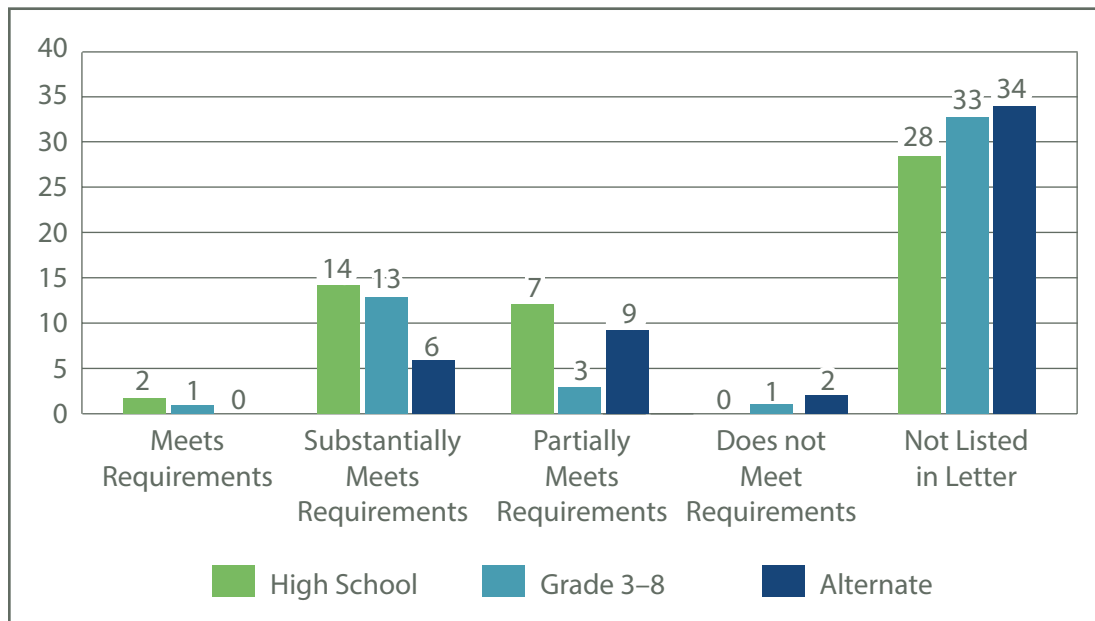
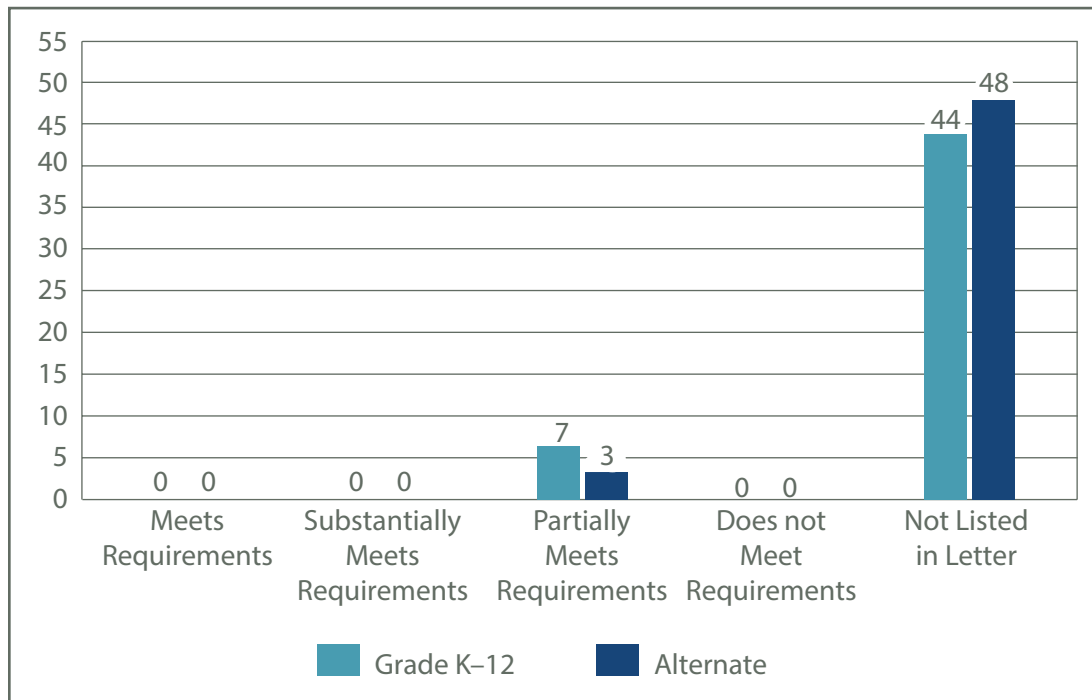


Chart 3. Peer Review Determinations for State English Language Proficiency Assessments



For a more detailed look at the peer review determinations for individual states and assessments, please refer to the [CSAI Critical Elements Analysis Chart](#), which presents the results for state decision letters and identifies which CEs states were requested to submit additional documentation for.

An analysis of the state decision letters indicates that a number of states were required to submit additional evidence, with some CEs addressed in multiple letters. The evidence requested for the same CE across decision letters sometimes differed. This report will analyze a sample of the CEs most widely cited as needing additional evidence submitted for general assessments in high school and grades 3–8 R/LA and mathematics, high school and grades 3–8 science assessments, alternate assessments across content areas and grade levels, and grades K–12 ELP assessments.

## High School R/LA and Mathematics

### CE 2.1 – Test Design and Development

- Appears in 28 decision letters for high school R/LA and mathematics assessments

CE 2.1 requires states to provide evidence that the state’s test design and test development process is well-suited for the content, is technically sound, aligns the assessments to the depth and breadth of the state’s academic content standards for the grade that is being assessed, and includes:

- Statement(s) of the purposes of the assessments and the intended interpretations and uses of the results;
- Test blueprints that describe the structure of each assessment in sufficient detail to support the development of assessments that are technically sound, measure the depth and breadth of the state’s grade-level academic content standards, and support the intended interpretations and uses of the results;
- Processes to ensure that each assessment is tailored to the knowledge and skills included in the state’s academic content standards, reflects appropriate inclusion of challenging content, and requires complex demonstrations or applications of knowledge and skills (i.e., higher-order thinking skills);

If the state administers a computer-adaptive assessment:

- The item pool and item selection procedures adequately support the test design and intended uses and interpretations of the results.
- It makes proficiency determinations with respect to the grade in which the student is enrolled and uses that determination for all reporting.

If the state administers a content assessment that includes portfolios, such an assessment may be partially administered through a portfolio but not entirely.<sup>3</sup>

The additional evidence requested to be submitted for CE 2.1 for high school R/LA and mathematics is varied, but the feedback provided to consortium states is generally consistent across decision letters.

## Peer Review Feedback

States administering a range of high school assessments for R/LA and mathematics received feedback to submit additional evidence that demonstrates how these assessments align with the depth and breadth of state academic content standards. Some states were also asked to submit a plan and timeline for addressing alignment issues previously identified in alignment studies ([AK, AR, AZ, CO, CT, DE, IA, IL, IN, KY, ME, MI, MN, MO, MS, MT, ND, NE, NH, OH, OK, RI, TN, UT, VA, WI, WV, WY decision letters](#)).

Some letters include examples of evidence that states may wish to consider submitting to meet the panel's requirements for this CE. For instance, [Kentucky's](#) decision letter included a suggestion for the state to conduct an independent alignment study that might provide evidence that test blueprints align to the full range of state academic content standards.

## Submission of Additional Evidence Requested

Based on ED's Map of the Critical Elements (see Appendix), CE 2.1 is likely addressed by coordinated evidence, so states in a consortium may wish to coordinate efforts to submit the additional evidence requested. However, states must also submit evidence specific to their state's standards/assessments.

Looking at the decision letters received by other states may provide some insight as to the types of additional evidence that states may choose to submit. A State's Guide also provides examples of evidence states may submit for CE 2.1 to address issues, though a state may determine that other types of evidence better address a CE.

A sampling of examples of evidence from A State's Guide that may address CE 2.1 includes:

- Relevant sections of state or regulations, language from contract(s) for the state's assessments, test coordinator or test administrator manuals, or other relevant documentation that clarifies the purposes of the assessment and the intended interpretations and uses of the results;
- Test blueprints that:
  - ◆ Describe the structure of each assessment in sufficient detail to support the development of a technically sound assessment, for example, in terms of the number of items, item types, the proportion of item types, response formats, the range of item difficulties, the types of scoring procedures, and applicable time limits;

<sup>3</sup>A State's Guide to the U.S. Department of Education's Assessment Peer Review Process. (2018). Washington, DC: U.S. Department of Education. Retrieved from <https://www2.ed.gov/admins/lead/account/saa/assessmentpeerreview.pdf>

- ◆ Align to the state’s grade-level academic content standards in terms of content (i.e., knowledge and cognitive process), the full range of the state’s grade-level academic content standards, the balance of content, and cognitive complexity;
- Documentation that the test design is tailored to the specific knowledge and skills in the state’s academic content standards (e.g., includes extended response items that require demonstration of writing skills if the state’s R/LA academic content standards include writing).
- Documentation of the approaches the state uses to include challenging content and complex demonstrations or applications of knowledge and skills (i.e., items that assess higher-order thinking skills, such as item types appropriate to the content that require synthesizing and evaluating information and analytical text-based writing or multiple steps and student explanations of their work); for example, this could include test specifications or test blueprints that require a certain portion of the total score be based on item types that require complex demonstrations or applications of knowledge and skills and the rationale for that design.

See *A State’s Guide* for a full list of potential examples of evidence that may address CE 2.1.

#### CE 5.4 – Monitoring Test Administration for Special Populations

- Appears in 21 decision letters for high school R/LA and mathematics assessments

CE 5.4 requires states to submit evidence that the state monitors test administration in its districts and schools to ensure that appropriate assessments, with or without appropriate accommodations, are selected for students with disabilities and English learners, so that they are appropriately included in assessments and receive accommodations that are:

- Consistent with the state’s policies for accommodation;
- Appropriate for addressing a student’s disability or language needs for each assessment administered;
- Consistent with accommodations provided to students during instruction and/or practice;
- Consistent with the assessment accommodations identified by a student’s individual educational program (IEP) team under IDEA, the placement team convened under Section 504; or for students covered by Title II of the Americans with Disabilities Act (ADA), the individual or team designated by a district to make these decisions; or another process for an English learner;
- Administered with fidelity to test administration procedures; and
- Monitored for administration of all required academic content assessments, alternate assessment based on alternate academic achievement standards (AA-AAAS), ELP assessments, and alternate ELP assessment.<sup>4</sup>

This CE will be checked for completeness by ED, and for consortium states, is likely to be addressed by a mix of state and coordinated evidence per the Map of the Critical Elements.

## Peer Review Feedback

States should be aware that the type of evidence submitted is important to the review process. Particularly for CEs related to assessment administration, monitoring, and training, documentation submitted to ED should include procedural evidence (e.g., documentation of procedures and protocols, studies that indicate how a particular feature or part of an assessment will function), as well as confirmatory evidence (e.g., proof that procedures and protocols were followed, analyses indicating that a particular feature or part of an assessment functioned as intended) ([AK, AZ, CA, CT, FL, GA, HI, IA, ID, IL, ME, MI, MS, MT, ND, NE, NH, RI, VA, WA, WY decision letters](#)).

<sup>4</sup> *A State’s Guide to the U.S. Department of Education’s Assessment Peer Review Process*. (2018). Washington, DC: U.S. Department of Education. Retrieved from <https://www2.ed.gov/admins/lead/account/saa/assessmentpeerreview.pdf>

In the case of CE 5.4, states may wish to submit procedural documentation regarding test administration (e.g., protocols, etc.), in addition to evidence showing that the prescribed monitoring protocols were followed during the assessment event.<sup>5</sup>

## Submission of Additional Evidence Requested

A State's Guide provides examples of evidence states may choose to submit to support their assertion that they meet CE 5.4.

Evidence to support this CE of the state's assessment system includes documents such as:

- Description of procedures the state uses to monitor that accommodations selected for students with disabilities, students covered by Section 504, and English learners are appropriate;
- Description of procedures the state uses to monitor that students with disabilities are placed by IEP teams in the appropriate assessment;
- The state's written procedures for monitoring the use of accommodations during test administration, such as guidance provided to districts; instructions and protocols for state, district, and school staff; and schedules for monitoring; and
- Summary of results of monitoring for the most recent year of test administration in the state.

The type(s) of documentation a state will choose to submit to fulfill the request for additional evidence should be guided by the specific feedback and comments provided to the state from peer reviewers and ED. Unlike CE 2.1 detailed above, the requests for additional evidence to be submitted for CE 5.4 cannot be compartmentalized by assessment system (i.e., ACT, PARCC, Smarter Balanced, or state-specific). States must look to their specific feedback to provide the documentation necessary to address the requests of the peer reviewers and ED. Examples of the types of documentation requested include:

- Evidence that test monitoring procedures and training address test accommodations and fidelity to test administration procedures ([Alabama decision letter](#));
- Evidence that the state monitors test administrations for special populations for consistency among individualized education programs, classroom instruction, and accommodations ([Rhode Island decision letter](#));
- Evidence that the state monitors test administration in districts and schools to ensure that appropriate assessments and accommodations are selected for students with disabilities under IDEA, students covered by Section 504, and English learners ([Iowa decision letter](#)).

## Grades 3–8 R/LA and Mathematics

CE 3.1 – Overall Validity, Including Validity Based on Content

- Appears in 24 decision letters for grades 3–8 R/LA and mathematics assessments

CE 3.1 asks states to submit evidence that shows adequate overall validity evidence for its assessments, which includes evidence that the state's assessments measure the knowledge and skills specified in the state's academic content standards, including:

<sup>5</sup>A State's Guide to the U.S. Department of Education's Assessment Peer Review Process. (2018). Washington, DC: U.S. Department of Education. Retrieved from <https://www2.ed.gov/admins/lead/account/saa/assessmentpeerreview.pdf>

- Documentation of adequate alignment between the state’s assessments and the academic content standards the assessments are designed to measure in terms of content (i.e., knowledge and process), balance of content, and cognitive complexity;
- Documentation that the assessments address the depth and breadth of the content standards; and
- If the state has adopted alternate academic achievement standards and administers alternate assessments aligned to those standards, the assessments show adequate alignment to the state’s academic content standards for the grade in which the student is enrolled in terms of content match (i.e., no unrelated content), and the breadth of content and cognitive complexity determined in test design are appropriate for students with the most significant cognitive disabilities.

This analysis focuses on validity evidence surrounding the assessments for grades 3–8 in R/LA and mathematics, not on validity evidence for a state’s alternate assessments based on alternate academic achievement standards.

On ED’s Map of the Critical Elements (see Appendix), CE 3.1 is shown as likely to be addressed by coordinated evidence.

## Peer Review Feedback

State decision letters included requests for states to submit evidence of alignment of grades 3–8 R/LA and mathematics assessments to state academic content standards, demonstrating that these assessments measure the breadth and depth of academic content standards, balance of content, appropriate inclusion of challenging content, and cognitive complexity. Some states were also asked to provide evidence of assessment usability, as it relates to assessment technology ([AK, AL, AR, AZ, FL, GA, IA, IN, KY, LA, MA, ME, MN, MS, ND, NH, OH, OK, RI, SC, TN, UT, WV, WY decision letters](#)).

North Dakota, a member of the Smarter Balanced consortium, was also asked to provide “evidence of a summary report that the CAT administered test forms matched test blueprints” ([North Dakota decision letter](#)).

## Submission of Additional Evidence Requested

States were asked to submit alignment evidence across grades 3, 4, 6, and 7 in both R/LA and mathematics. Submission of alignment evidence at grades 5 and 8 is not sufficient for meeting ED’s requirements for CE 3.1, and states will have to go back to their respective test developers and request that they compile the stated alignment evidence. Many statements of additional evidence requested are tied to ED’s requests for additional evidence of CE 2.1, as well as evidence of alignment or improved alignment based upon previously submitted evidence.

### CE 6.4 – Reporting

- Appears in 23 decision letters for grades 3–8 R/LA and mathematics assessments

CE 6.4 asks states to submit evidence that the state reports its assessment results for all assessed students, and the reporting facilitates timely, appropriate, credible, and defensible interpretations and uses of the results (for students tested) by parents, educators, state officials, policymakers and other stakeholders, and the public. The state must report on student academic achievement for all students and each student group at each achievement level.



For academic content assessment, the state reports assessment results, including itemized score analyses, to districts and schools so that parents, teachers, principals, and administrators can interpret the results and address the specific academic needs of students, and the state also provides interpretive guides to support appropriate uses of the assessment results.

- The state provides for the production and delivery of individual student interpretive, descriptive, and diagnostic reports after each administration of its academic content assessments that:
  - ◆ Provide valid and reliable information regarding a student’s academic achievement;
  - ◆ Report the student’s academic achievement in terms of the state’s grade-level academic achievement standards;
  - ◆ Provide information to help parents, teachers, and principals interpret the test results and address the specific academic needs of students;
  - ◆ Are provided in an understandable and uniform format;
  - ◆ Are, to the extent practicable, written in a language that parents and guardians can understand or, if it is not practicable to provide written translations to a parent or guardian with limited English proficiency, are orally translated for such a parent or guardian; and
  - ◆ Upon request by a parent who is an individual with a disability as defined by the ADA, as amended, are provided in an alternative format accessible to that parent.
- The state follows a process and timeline for delivering individual student reports to parents, teachers, and principals as soon as is practicable after each test administration.

It has been indicated on the Map of the Critical Elements that CE 6.4 is likely addressed with a mix of coordinated (for consortium states) and state-specific evidence, and will also be checked for completeness by ED.

## Peer Review Feedback

For this CE, states were asked to submit evidence showing that individual student score reports are available in alternate formats (upon request), as well as documentation regarding the process and timeline for delivery of individual score reports to parents, teachers, and principals. States were also asked for documentation of score reports that facilitate itemized score analyses, as well as ensuring that score reports include data on student achievement relative to the state’s grade-level academic content standards ([AK, AL, AR, CA, FL, ID, LA, MA, ME, MN, MS, NE, NH, OH, OK, OR, RI, SC, TN, VA, VT, WA, WY decision letters](#)).

## Submission of Additional Evidence Requested

A *State’s Guide* explains that the evidence and supporting documentation states choose to submit to ED for review of CE 6.4 must demonstrate that the state’s reporting system facilitates timely, appropriate, credible, and defensible interpretations and uses of its assessment results. However, states must also protect the privacy of test takers when this information is submitted. Samples should be redacted to protect personally identifiable information or populated with information about a fictitious student for illustrative purposes. To support this CE, a state may provide the same types of evidence for grades 3–8 and high school general assessments, as well as for alternate assessments administered (if required based on feedback).

## High School Science Assessments

### CE 3.1 – Overall Validity, Including Validity Based on Content

- Appears in 14 decision letters for high school science assessments

CE 3.1 asks states to submit evidence that shows adequate overall validity evidence for its assessments, which includes evidence that the state’s assessments measure the knowledge and skills specified in the state’s academic content standards, including:

- Documentation of adequate alignment between the state’s assessments and the academic content standards the assessments are designed to measure in terms of content (i.e., knowledge and process), balance of content, and cognitive complexity;
- Documentation that the assessments address the depth and breadth of the content standards; and
- If the state has adopted alternate academic achievement standards and administers alternate assessments aligned to those standards, the assessments show adequate alignment to the state’s academic content standards for the grade in which the student is enrolled in terms of content match (i.e., no unrelated content) and the breadth of content and cognitive complexity determined in test design are appropriate for students with the most significant cognitive disabilities.

On ED’s Map of the Critical Elements (see Appendix), CE 3.1 is shown as likely to be addressed by coordinated evidence.

## Peer Review Feedback

In the decision letters for reviewed high school science assessments, states were asked to provide evidence of alignment between assessments and state science standards ([AR](#), [CO](#), [DE](#), [GA](#), [IA](#), [ID](#), [IN](#), [KY](#), [NE](#), [OH](#), [UT](#), [VA](#), [WV](#), [WY](#) [decision letters](#)). This may include findings from independent studies of statewide assessments, content standards, and achievement standards. Some states were also asked to submit a plan and timeline for improving the alignment between assessments and state science standards.

## Submission of Additional Evidence Requested

To address alignment between state content standards and assessments, states might submit results from an independent alignment study that is technically sound (i.e., method and process, appropriate units of analysis, clear criteria) and documents the following:

- Each assessment is aligned to its test blueprint and each blueprint addresses the depth and breadth of the state’s academic content standards; and
- The state follows procedures to ensure alignment during test development.

States may also include a description of a systematic process and timeline they will implement to address any gaps or weaknesses identified in the alignment studies.

## Grades 3–8 Science Assessments

### CE 3.1 – Overall Validity, Including Validity Based on Content

- Appears in 11 decision letters for grades 3–8 science assessments

CE 3.1 asks states to submit evidence that shows adequate overall validity evidence for its assessments, which includes evidence that the state’s assessments measure the knowledge and skills specified in the state’s academic content standards, including:

- Documentation of adequate alignment between the state’s assessments and the academic content standards the assessments are designed to measure in terms of content (i.e., knowledge and process), balance of content, and cognitive complexity;
- Documentation that the assessments address the depth and breadth of the content standards; and
- If the state has adopted alternate academic achievement standards and administers alternate assessments aligned to those standards, the assessments show adequate alignment to the state’s academic content standards for the grade in which the student is enrolled in terms of content match (i.e., no unrelated content) and the breadth of content and cognitive complexity determined in test design are appropriate for students with the most significant cognitive disabilities.

On ED’s Map of the Critical Elements (see Appendix), CE 3.1 is shown as likely to be addressed by coordinated evidence.

## Peer Review Feedback

In the decision letters, states were requested to provide documentation of the alignment between state assessment and science standards, as well as evidence of adequate validity based on test content ([AL, AR, CO, DE, GA, IA, IL, MN, MO, OH, UT, VA decision letters](#)). Some states were also asked to submit evidence that instances of weak alignment between state assessments and science standards were addressed.

## Submission of Additional Evidence Requested

To address the alignment between state content standards and assessments, states might submit results from an independent alignment study that is technically sound (i.e., method and process, appropriate units of analysis, clear criteria) and documents the following:

- Each assessment is aligned to its test blueprint and each blueprint addresses the depth and breadth of the state’s academic content standards; and
- The state follows procedures to ensure alignment during test development.

States may also include a description of a systematic process and timeline they will implement to address any gaps or weaknesses identified in the alignment studies.

To address assessment validity based on assessment content, states might submit evidence that demonstrates how the assessment’s levels of validity are generally consistent with expectations of current professional standards, such as:

- Test blueprints for the overall assessment and/or all domain-specific sub-tests, as submitted under CE 2.1;
- Logical or empirical analyses that show the assessment content adequately represents the depth and breadth of state academic content standards;

- Report of expert judgement of the relationship between assessment components and the state’s academic content standards; or
- For academic content standards, reports of analyses to demonstrate that the state’s assessment of academic content is appropriately related to the specific inferences made from assessment scores about student proficiency in the state’s academic content standards for all student groups.

### Alternate Assessments (R/LA, Mathematics, Science)

Some states were required to submit additional evidence for CEs across different combinations of assessment types within their assessment systems. CE 4.1 and CE 5.1 were chosen to be examined specifically for the Alternate Assessment portion of this brief because they had some of the highest occurrences of needing additional evidence submitted for a state’s alternate assessment suite. States that were given feedback with regard to their “entire assessment system” for a particular CE have been included.

#### CE 4.1 – Reliability

- Appears in 27 decision letters for Alternate Assessments

CE 4.1 requires that the state has documented adequate reliability evidence for its assessments for the following measures of reliability for the state’s student population overall and each student group, consistent with nationally recognized professional and technical testing standards. If the state’s assessments are implemented in multiple states, measures of reliability for the assessment overall and each student group are consistent with nationally recognized professional and technical testing standards, including:

- Test reliability of the state’s assessments estimated for its student population;
- Overall and conditional standard error of measurement of the state’s assessments, including any domain or component sub-tests, as applicable;
- Consistency and accuracy of estimates in categorical classification decisions for the cut scores, achievement levels, or proficiency levels based on the assessment results; and
- For computer-adaptive tests, evidence that the assessments produce test forms with adequately precise estimates of a student’s academic achievement.

Coordinated evidence likely addresses CE 4.1 for consortium states.

Though use of the word “adequate” in the description of CE 4.1 does not provide an explicit level of detail on the volume of evidence that will meet ED’s requirements, A State’s Guide provides states with examples of evidence they may wish to provide.

### Peer Review Feedback

The decision letters sent to states that contained requests for additional documentation and evidence for CE 4.1 for alternate assessments included feedback across R/LA and mathematics content areas for all submitted grade levels. Not all states are required to submit additional evidence for all subject areas/grade levels for their alternate assessments. Decision letters indicate for which components of a state’s alternate assessment suite additional evidence is required.

States that submitted evidence for peer review, administer the National Center and State Collaborative Multi-State Alternate Assessment (NCSC/MSAA), and are requested to submit additional evidence for CE 4.1 were provided with the same feedback, directing them back to the explanatory note(s) already provided for writing items in CE 3.1 – Overall Validity, including Validity Based on Content. CE 3.1 has been analyzed above in the context of a state’s grades 3–8 general assessments, but many of the same types of evidence submitted to address CE 3.1 for the NCSC/MSAA can also be used to provide additional evidence to support a state’s claim that it meets requirements for CE 4.1 ([AR, AZ, DC, MD, ME, MT, NM, RI, SD decision letters](#)).

States that administer the Dynamic Learning Maps Integrated (DLM-Int) assessment model for their alternate R/LA and mathematics assessments are also directed back to the feedback provided for CE 3.1 and are asked specifically for “evidence of model fit analysis.” These states are also required to submit “evidence regarding consistency and accuracy of classifications, which will be satisfied by the response to element 3.3.” CE 3.3 – Validity Based on Internal Structure is likely addressed by coordinated evidence, and requires that a state “has documented adequate validity evidence that the scoring and reporting structures of its assessments are consistent with the sub-domain structures of the state’s academic content standards on which the intended interpretations and uses of results are based” ([IA, VT decision letters](#)).

States administering the Dynamic Learning Maps Year End (DLM-YE) model for R/LA and mathematics are required to submit “evidence of monitoring and refinement of the diagnostic classification models from subsequent test administrations” ([IL, WI decision letters](#)).

For alternate assessments that require additional documentation to meet the requirements of CE 4.1, states must submit a variety of documentation across content areas and grade levels (available in ED’s comments to the state in the decision letter issued). The documentation requested includes “evidence of... reliability estimates by gender and dichotomous demographic categories to demonstrate lack of bias” ([Delaware decision letter](#)) and “evidence of... inter-rater reliability” ([Wyoming decision letter](#)), among others.

#### CE 5.1 – Procedures for Including Students with Disabilities

- Appears in 27 decision letters for alternate assessments

CE 5.1 requires states to submit evidence that it has in place procedures to ensure the inclusion of all public elementary and secondary school students with disabilities in the state’s assessment system. Decisions about how to assess students with disabilities must be made by a student’s IEP team under IDEA, the placement team under Section 504, or the individual or team designated by a district to make that decision under Title II of the ADA, as applicable, based on each student’s individual abilities and needs.

If a state adopts alternate academic achievement standards for students with the most significant cognitive disabilities and administers an alternate assessment aligned to those standards under ESEA section 1111(b)(1)(E) and (b)(2)(D), respectively, the state must:

- Establish guidelines for determining whether to assess a student with an AA-AAAS, including:
  - ◆ A state definition of “students with the most significant cognitive disabilities” that addresses factors related to cognitive functioning and adaptive behavior;

- Provide information to IEP teams to inform decisions about student assessments that:
  - ◆ Provides a clear explanation of the differences between assessments aligned to grade-level academic achievement standards and those aligned to alternate academic standards, including any effects of state and local policies on a student's education resulting from taking an AA-AAAS, such as how participation in such assessments may delay or otherwise affect the student from completing the requirements for a regular high school diploma;
- Ensure that parents of students assessed with an AA-AAAS are informed that their child's achievement will be measured based on alternate academic achievement standards;
- Not preclude a student with the most significant cognitive disabilities who takes an AA-AAAS from attempting to complete the requirements for a regular high school diploma;
- Promote, consistent with the requirements under IDEA, the involvement and progress of students with the most significant cognitive disabilities in the general education curriculum that is based on the state's academic content standards for the grade in which the student is enrolled;
- Develop, disseminate information on, and promote the use of appropriate accommodations to ensure that a student with the most significant cognitive disabilities who does not take an AA-AAAS participates in academic instruction and assessments for the grade in which the student is enrolled; and
- Have in place and monitor implementation of guidelines for IEP teams to apply in determining, on a case-by-case basis, which students with the most significant cognitive disabilities will be assessed based on alternate academic achievement standards, if applicable. Such guidelines must be developed in accordance with 34 CFR § 200.6(d).

CE 5.1 contains many components, and sufficient evidence must be submitted for each to meet the peer review requirements and comply with ESSA.

## Peer Review Feedback

The additional evidence these states are required to submit includes:

- Evidence including guidelines for IEP teams to use in determining whether to assess a student on the general assessment without accommodation(s), the general assessment with accommodation(s), or an alternate assessment ([IA, KY decision letters](#));
- Evidence that the state provides clear explanations of the differences between assessments based on grade-level academic achievement standards and assessments based on alternate achievement standards ([KY, WI decision letters](#)); and
- Evidence that sufficient information is provided to parents of students with the most significant cognitive disabilities, and that they are informed their student's achievement will be based on alternate academic achievement standards and informed of any possible consequences of taking the AA-AAAS resulting from district or state policy (e.g., ineligibility for a regular high school diploma if the student does not demonstrate proficiency in the content area on the state's general assessments) ([OR, TX, VT, WY decision letters](#)).

The additional evidence requested is not delineated by assessment type. States that administer the same alternate assessments (e.g., DLM-YE, DLM-INT, or NCSC/MSAA) may be required to submit different pieces of additional evidence, most likely because much of the documentation needed to meet the requirements of CE 5.1 consists of communication and protocols at the state and district levels. Additionally, the Map of the Critical Elements indicates that CE 5.1 is "likely addressed by state-specific evidence."

## Submission of Additional Evidence Requested

Examples of evidence to support CE 5.1 are provided in A State's Guide. However, based on the feedback they receive from ED, states may have to submit additional documentation not listed in the examples. This may include evidence of stakeholder communication that outlines the key differences between assessments based on grade-level academic standards and those based on alternate achievement standards, documentation of evidence provided to parents of students with the most significant cognitive disabilities, and evidence of the procedures and guidelines in place for use by students' individualized education program (IEP) teams to determine which assessment/accommodations are needed by each student.

## English Language Proficiency Assessments

In April 2019, ED conducted the first peer review of state ELP assessments; all states were expected to submit documentation of their ELP assessments at this time. As of August 2019, seven states have received decision letters on their state's ELP assessment: [AK](#), [AL](#), [AR](#), [DE](#), [IA](#), [MD](#), and [NY](#). All seven states received *partially meets requirements* determinations. Common across these seven states is the request to submit additional documentation for the following critical elements:

- 2.1 – Test Design and Development
  - ◆ Evidence of test blueprints that describe the structure of state ELP assessments.
  - ◆ Evidence that state ELP assessments are aligned to the depth and breadth of state ELP standards.
  - ◆ Evidence that the item pool and item selection procedures adequately support multi-stage adaptive administrations.
  - ◆ Evidence that proficiency determinations are made with respect to the grade a student is enrolled in.
- 2.2 – Item Development
  - ◆ Evidence of reasonable and technically sound procedures to develop and select items (e.g., timeline of development, qualifications of item writers, item-writing training, item review processes and reviewer qualifications, field test processes for each domain, and technical advisory committee review).
- 3.1 – Overall Validity, Including Validity Based on Content
  - ◆ Documentation of alignment between a state's ELP assessment and ELP standards, in terms of language, knowledge, and skills and the depth and breadth of ELP standards.
  - ◆ Documentation of alignment between state ELP standards and the language demands of state academic standards.
- 3.2 – Validity Based on Cognitive Processes
  - ◆ Evidence that state ELP assessments tap intended language processes appropriate for the grade levels and grade bands represented in state ELP standards.
- 3.3 – Validity Based on Internal Structure
  - ◆ Evidence that ELP assessment scoring and reporting structures are consistent with the sub-domain structures of state ELP standards.

- 4.1 – Reliability
  - ◆ Evidence of test reliability, including:
    - Reliability by subgroups.
    - Consistency and accuracy of estimates in categorical classification decisions for the cut scores, achievement levels, or proficiency levels based on assessment results.
    - Reliability statistics used to inform ongoing maintenance and development.
  - ◆ Evidence that assessments provide test forms with adequately precise estimates of English proficiency.
- 4.2 – Fairness & Accessibility
  - ◆ Evidence that the state ELP assessments are accessible to all students and subgroups in terms of design, development, and analysis (e.g., implementation of universal design principles, evidence of item writer training materials that address accessibility, qualifications of item writers and item reviewers).
- 4.3 – Full Performance Continuum
  - ◆ Evidence that each assessment provides an adequately precise estimate of student performance across the full performance continuum for ELP assessments.
- 4.4 – Scoring
  - ◆ Evidence of standardized scoring procedures and protocols that are designed to produce reliable and meaningful results, facilitate score interpretations, and report assessment results in terms of state ELP standards.
  - ◆ Evidence that if an English learner has a disability that precludes assessment of the student in one or more of the required domains/components (listening, speaking, reading, and writing) because there are no appropriate accommodations for the affected domain(s)/component(s), the state ensures that the student is assessed in the remaining domain(s)/component(s), including a description of how this will occur.
- 5.3 – Accommodations
  - ◆ Evidence that the provided accommodations:
    - Are appropriate and effective for meeting the individual student’s need(s) to participate in the assessments.
    - Do not alter the construct being assessed.
    - Allow meaningful interpretations of the results and comparison of scores for students who need and receive accommodations and students who do not need and do not receive accommodations.
  - ◆ Evidence of a process to individually review and allow exceptional requests for a small number of students who require accommodations beyond those routinely allowed.
- 6.4 – Reporting
  - ◆ Evidence that the state’s reporting of assessment results facilitates timely interpretations and uses of those results by parents, educators, state officials, policymakers and other stakeholders, and the public.
  - ◆ Evidence that the state provides coherent and timely information about each student’s attainment of the state’s ELP standards to parents that are, to the extent practicable, written in a language that parents and guardians can understand or, if it is not practicable to provide written translations to a parent or guardian with limited English proficiency, are orally translated.
  - ◆ Evidence that student reports are, upon request by a parent with a disability, provided in an alternative format accessible to that parent.



- ◆ Evidence that performance-level descriptors are included in student score reports.
- ◆ Evidence that the state reports to the public results of the ELP of all English learners.

Only three states ([AK](#), [DE](#), [MD](#)) have received a decision letter based on a peer review of their alternate ELP assessment for grades K–12. All three states received *partially meets requirements* determinations for their state alternate ELP assessments and were asked to submit additional evidence for the following CEs:

- 1.3 – Required Assessments
  - ◆ Evidence that the alternate ELP assessment is available in kindergarten.
- 1.4 – Policies for Including All Students in Assessments
  - ◆ Evidence that the alternate ELP assessment is available in kindergarten.
- 2.1 – Test Design & Development
  - ◆ Evidence that assessments are aligned to the depth and breadth of the state’s ELP standards, including:
    - Statement of the purposes and intended uses of the results.
    - Test blueprints.
    - Processes to ensure that the ELP assessment is tailored to the knowledge and skills included in the state’s ELP standards and reflects appropriate inclusion of the range of complexity found in the standards (e.g., detail about the routing rules, detail of the item selection process for paper forms to ensure it adheres to the blueprint).
- 2.2 – Item Development
  - ◆ Evidence of reasonable and technically sound procedures to develop and select items to assess ELP (e.g., involvement of experts with knowledge of English Learners with significantly cognitive disabilities).
- 2.3 – Test Administration
  - ◆ Evidence of established communication to educators of clear, thorough, and consistent standardized procedures for the administration of its assessments, including administration with accommodations (e.g., guidelines or recommended qualifications of test administrators, including volunteers, training of volunteers, and qualifications and training for the human providers of accommodations).
  - ◆ Evidence of established procedures to ensure that general and special education teachers, paraprofessionals, teachers of English Learners, specialized instruction support personnel, and other appropriate staff receive the necessary training to know how to administer assessments, including, as necessary, alternate assessments, and know how to make use of appropriate accommodations during assessments for all students with disabilities (e.g., content of training modules, evidence that training is required for test administrators, and evidence of participation in such training).
- 2.5 – Test Security
  - ◆ Evidence of detection of test irregularities.
  - ◆ Evidence of remediation following any test security incidents.
  - ◆ Evidence of the investigation of alleged or factual test irregularities (e.g., forensic analysis and plans to address test security concerns).

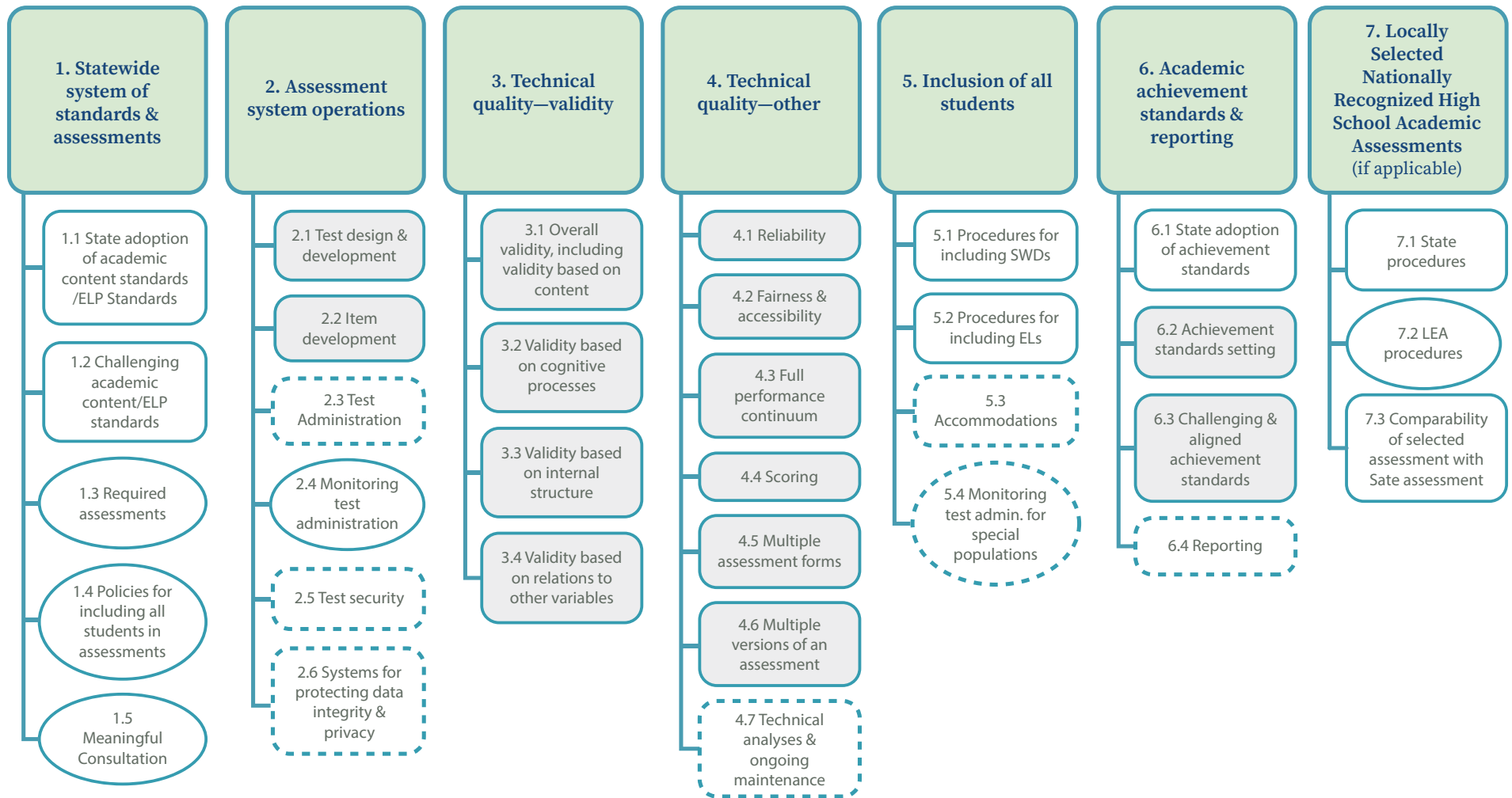
- ◆ Evidence of policies and procedures that prevent assessment irregularities, including maintaining the security of test materials (both during test development and at time of test administration), proper test preparation guidelines and administration procedures, incident-reporting procedures, consequences for confirmed violations of test security, and requirements for annual training at the district and school levels for all individuals involved in test administration, including:
  - Evidence of communication that the state’s security policies include ACCESS and Alternate ACCESS.
  - Evidence of policies and procedures for the Alternate ACCESS to protect the integrity of the test given and evidence that the test form has been unchanged for the past several years.
- 3.1 – Overall Validity, Including Validity Based on Content
  - ◆ Evidence of adequate linkage to the state’s ELP standards in terms of content match (i.e., no unrelated content) and that the breadth of content and linguistic complexity determined in test design are appropriate for English Learners with the most significant cognitive disabilities.
- 3.2 – Validity Based on Cognitive Processes
  - ◆ Adequate validity evidence that the assessments tap the intended language processes appropriate for each grade level/grade band as represented in the state’s ELP standards.
- 3.3 – Validity Based on Internal Structure
  - ◆ Evidence that the scoring and reporting structures of the assessments are consistent with the sub-domain structures of the state’s ELP standards (e.g., an explanation of how the included statistical analyses relate to the validity framework for the assessments).
- 3.4 – Validity Based on Relations to Other Variables
  - ◆ Adequate validity evidence that the state’s assessment scores are as expected with other variables (e.g., interpretations for studies of relationships between the state academic assessments and the ELP assessments).
- 4.1 – Reliability
  - ◆ Evidence of test reliability, including:
    - Reliability by subgroups.
    - Consistency and accuracy of estimates in categorical classification decisions for the cut scores, achievement levels, or proficiency levels based on the assessment results.
    - Evidence that reliability statistics are used to inform ongoing maintenance and development.
  - ◆ Evidence of reliability, including test information functions for overall composite scores.
- 4.2 – Fairness & Accessibility
  - ◆ Evidence that the assessments are accessible to all students and fair across student groups in design, development, and analysis (e.g., the implementation of universal design principles, to the extent practicable, during item development and review, and additional differential item functioning analyses to include more student subgroups).
  - ◆ Evidence that the state has taken reasonable and appropriate steps to ensure that its assessments are accessible to all EL students and fair across student groups, including English Learners with disabilities, in their design, development, and analyses, as well as guidance and instructions on appropriate instructional supports that can be used during the assessment, particularly for braille and alternate modes of communication.

- 4.3 – Full Performance Continuum
  - ◆ Evidence that each assessment provides an adequately precise estimate of student performance across the full performance continuum for ELP assessments, including performance for EL students with high and low levels of ELP.
- 4.4 – Scoring
  - ◆ Evidence that if an English Learner has a disability that precludes assessment of the student in one or more of the required domains/components (listening, speaking, reading, and writing) such that there are no appropriate accommodations for the affected domain(s)/component(s), the state must provide a description of how it will ensure that the student is assessed in the remaining domain(s)/component(s) in which it is possible to assess the student, and a description of how this will occur.
  - ◆ Evidence of the implementation of standardized scoring procedures and protocols, including definitions of key terms and test administration and scoring procedures.
- 4.5 – Multiple Assessment Forms
  - ◆ Evidence that all forms adequately represent the state’s ELP standards and yield consistent score interpretations, such that the forms are comparable within and across settings (e.g., evidence that using the same test items every year does not impact validity).
- 4.7 – Technical Analyses & Ongoing Maintenance
  - ◆ Evidence of adequate technical quality is made public, including on the state’s website.
  - ◆ Evidence of a system for monitoring, maintaining, and improving, as needed, the quality of its assessment system.
- 5.1 – Procedures for Including Students with Disabilities
  - ◆ Evidence of policies that require the inclusion of an English Learner with a disability that precludes assessment of the student in one or more of the required domains (i.e., ensuring that the student will be assessed based on the remaining component(s) in which it is possible to assess the student).
- 5.3 – Accommodations
  - ◆ Evidence that the provided accommodations:
    - Are appropriate and effective for meeting the individual student’s need(s) to participate in the assessments.
    - Do not alter the construct being assessed.
    - Allow meaningful interpretations of the results and comparison of scores for students who need and receive accommodations and students who do not need and do not receive accommodations.
  - ◆ Evidence of a process to individually review and allow exceptional requests for a small number of students who require accommodations beyond those routinely allowed.
  - ◆ Evidence that appropriate accommodations are available for English Learners.
  - ◆ Evidence that accommodations do not deny students with disabilities or English Learners the opportunity to participate in the assessment.
- 6.1 – State Adoption of Achievement Standards
  - ◆ Evidence that the state adopted ELP achievement standards that address the different proficiency levels of English Learners.
  - ◆ If the state has developed alternate ELP achievement standards, evidence that it has adopted them only for English Learners with the most significant cognitive disabilities who cannot participate in the regular ELP assessment even with appropriate accommodations.

- 6.2 – Achievement Standards Setting
  - ◆ Evidence that the state used a technically sound method and process for setting ELP achievement standards, such that cut scores are developed for every grade level/grade band, content domain/language domain, and/or composite for which proficiency-level scores are reported.
- 6.3 – Challenging & Aligned Achievement Standards
  - ◆ If the state has developed alternate ELP achievement standards, evidence that they are linked to the state's grade-level/grade-band ELP standards and reflect professional judgment of the highest ELP achievement standards possible for English Learners with the most significant cognitive disabilities.
- 6.4 – Reporting
  - ◆ Evidence that the state's reporting of assessment results facilitates timely interpretations and uses of those results by parents, educators, state officials, policymakers and other stakeholders, and the public.
  - ◆ Evidence that the state provides coherent and timely information about each student's attainment of the state's ELP standards to parents that are, to the extent practicable, written in a language that parents and guardians can understand or, if it is not practicable to provide written translations to a parent or guardian with limited English proficiency, are orally translated for such a parent or guardian.
  - ◆ Evidence that student reports are, upon request by a parent with a disability, provided in an alternative format accessible to that parent.
  - ◆ Evidence that performance-level descriptors are included on student score reports.

Though not all states have received decision letters from ED, these letters might provide insights into peer reviewer feedback related to states' ELP and alternate ELP assessments.

### Appendix: Map of the Critical Elements for the State Assessment System Peer Review



**KEY**

- Critical elements in ovals will be checked for completeness by Department staff; if necessary, they may also be reviewed by assessment peer reviewers (e.g. Critical Element 1.3). All other critical elements will be reviewed by assessment peer reviewers.
- Critical elements in clear boxes with solid outlines likely will be addressed with State-specific evidence, even if a State administers the same assessments administered by other States (e.g. Critical Element 5.1).
- Critical elements in ovals or clear boxes with dashed outlines likely will be addressed by both State-specific evidence and coordinated evidence for States administering the same assessments (e.g. Critical Element 2.3, 5.4).
- Critical elements in shaded boxes likely will be addressed by coordinated evidence for all States administering the same assessments (e.g. Critical Element 2.1).